Bandolier 140

Independent evidence-based thinking about health care

On evidence-based policy making

Yes, many of you cynics out there will be thinking that this is a bit of an oxymoron, combining incongruous and contradictory terms. Evidence-based policy making must exist: it is just that we never see the processes. Yet we know that processes exist to help, like health technology assessment, or NICE in the UK and similar organisations in other countries. Inevitably it all gets a bit bureaucratic, and decision-making can seem glacial, but that is the price we pay for doing things right.

The long view is that this must be the right way of doing things. Yes there are glitches and delays, but if we keep grinding away at it we'll get it right in the end, and healthcare will enter the broad, sunny, uplands we have been waiting for.

Policy-based evidence

There is an alternate view though, or possibly views. Two main objections to evidence-based policy making argue that it will never work.

The first is that top-down policy will always fail. Professionals resent it, it will often be wrong or at best out of date, the bureaucracy will subvert it to its own ends, and the result will be order, counter-order, and disorder. Some would say that is about where we are in the UK right now.

The other argument is that large organisations run on budgets deploy any arguments they like to make the evidence fit the budget. It can all get very murky, and because professionals are not fools and can spot policy-based evidence, it makes for a generally uncomfortable time all round.

Alternatives

"Trust the professionals" might be a slogan for at least one alternative, and give them the tools to finish the job. It means letting small groups who have a sense of ownership for one part of the whole get on with their own bottom up redesign, and develop care pathways based on best evidence and local circumstance.

This issue of Bandolier has a lovely example of improved diabetes care from a large care pathway in Israel. Using good IT, best evidence, and multidisciplinary ownership, they created a system that did better for patients, gave more satisfaction to professionals, and at a lower cost per patient. Success.

Treating thrush

Most women of reproductive years experience at least one episode of vulvovaginal candidiasis (thrush), caused by infection with one or more species of Candida, most often albica. Anti-fungal drugs can be administered orally or vaginally. About 1 in 10 women who experience an episode of thrush can go on to develop recurrent candidiasis, though this has no recognised risk factors. Management of recurrent candidiasis is difficult, and Bandolier readers have asked for evidence. A Cochrane review [1] has looked at oral versus intra-vaginal anti-fungal treatments of uncomplicated thrush, and a recent randomised trial [2] examined treating recurrent infection.

Uncomplicated thrush

The Cochrane review [1] examined uncomplicated thrush, acute episodes occurring less frequently than four times a year in women aged 16 years or older. Diagnosis was by culture or microscopy, and studies with immunocompromised, pregnant, breast feeding or diabetic women were not included.

Trials had to be randomised, and compare any imidazole or triazole anti-fungal used vaginally with an oral equivalent (fluconazole or itraconazole). Treatments generally lasted less than a week. Various outcomes were examined, including clinical cure in the short term (usually about one or two weeks) or long term (generally about four weeks).

Results

There was no difference in long term clinical or mycological cure for oral fluconazole compared with vaginal clotrimazole. Seven trials with 836 women had long term cure rates of 83% for oral and 82% for intravaginal treatment (Figure 1), as well as high rates of women with mycological cure on culture or microscopy.

In this issue

Figure 1: Long-term cure with oral or intra-vaginal imidazole or triazole antifungal

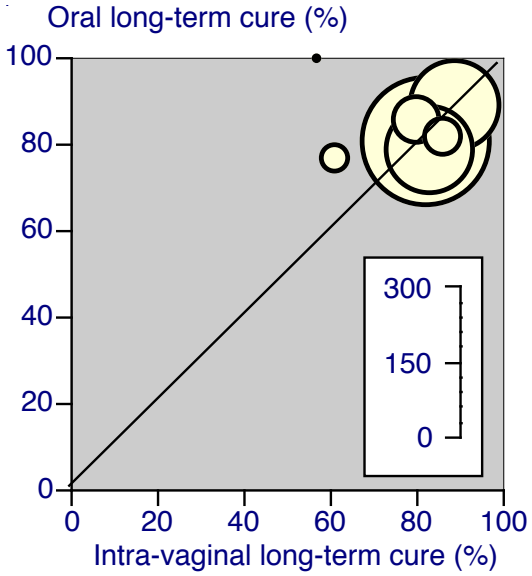
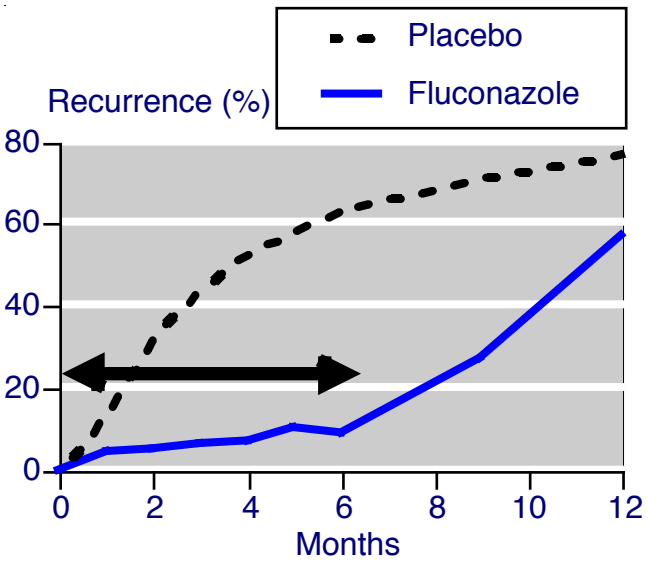


Figure 2: Recurrence of thrush with weekly fluconazole or placebo for six months (arrow)



Recurrent thrush

In a large randomised trial [2], women aged 18 years or older were required to have active candida vaginitis, with at least four documented episodes in the previous 12 months, and positive culture or microscopy of vaginal secretions. Clinical scoring was based on presence of symptoms of pruritus, burning, or irritation, and signs of erythema, oedema, and excoriation or fissures (each scored 0-3, maximum score 18). Women excluded were those with negative culture, who were pregnant, had mixed infections, had previous recent anti-fungal treatment, or who were immunocompromised.

There was an induction phase to ensure that women fulfilled entry criteria, followed by receipt of three 150 mg oral doses of fluconazole over nine days. At 14 days women had a vaginal examination and were entered into the trial if they had a negative culture and a clinical cure (symptom score of 3 or less out of 18).

Treatment was with a single 150 mg oral dose of fluconazole or placebo tablet every week for six months. Clinic visits occurred every month for six months, then at nine and 12 months. Clinical scoring and detailed pelvic examinations for fungal culture were made at these visits.

Results

The main analysis was on 343 women initially clear of thrush, with an average age of 34 (range 18 to 65) years. While all were cured at the start of the study, recurrence was rapid with placebo treatment (Figure 2), so that almost half had a recurrence by three months, two-thirds by six months, and four out of 10 by 12 months.

Weekly treatment with oral fluconazole meant that over the six-month treatment period, only 1 in 10 women had a recurrence of thrush (Figure 2). When treatment stopped, however, there was a rapid increase in recurrence, so that about half of the women had a recurrence after a further six months without fluconazole treatment.

At the end of six months of treatment of women with recurrent thrush initially free from symptoms and with negative vaginal culture, 9% of women had recurrence with fluconazole compared with 64% with placebo. The number needed to treat for six months with weekly oral fluconazole 150 mg for one woman continuing to be free of thrush was 1.8 (1.6 to 2.2). There was no indication of the emergence of resistant strains of Candida.

Adverse event

Adverse events with these treatments appear to be few. The review of oral and vaginal anti-fungal treatments [1], and the randomised trial [2] reported two withdrawals because of adverse events from treatment with oral fluconazole, or other oral anti-fungals.

Comment

Oral antifungal agents are highly effective for treating a single episode of thrush, and oral fluconazole is available without prescription in some countries. Weekly oral fluconazole is effective for recurrent thrush, though not after treatment stops, even for six months.

Weekly treatment is expensive for health services or individuals, costing about £350 a year. Whether weekly treatment is better than twice weekly or monthly oral fluconazole is not known. We don't know the optimal strategy, but long-term cure remains elusive.

References:

1 MC Watson et al. Oral versus intravaginal imidazole and triazole anti-fungal treatment of uncomplicated vulvovaginal candidiasis (thrush). The Cochrane Database of Systematic Reviews 2001 issue 3.
2 JD Sobel et al. Maintenance fluconazole therapy for recurrent vulvovaginal candidiasis. New England Journal of Medicine 2004 351: 876-883.

Circumcision for UTI

Circumcision is the most common surgery carried out on children, for various reasons. Rates vary because of cultural and religious differences. Circumcision rates are as high as two-thirds in North America, 10-20% in Australia, but much lower in Europe and much of Asia. Circumcision is recommended for true phimosis where the foreskin cannot be drawn back to uncover the glans penis, or inflammation of the glans penis, or for recurrent urinary tract infection (UTI). How effective circumcision is in reducing UTI a systematic review [1] tells us.

Systematic review

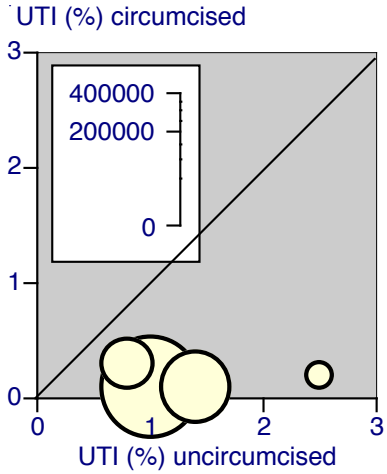
The review sought studies of any design to the end of 2002, not limited by language, in several electronic databases and bibliographies. Studies with information on the effect of male circumcision on UTI were included if they provided enough information to be able to compare odds of UTI in a circumcised group with the odds in an uncircumcised group. There was no age restriction.

Results

Twelve studies fulfilled the inclusion criteria. Most examined infants, though one included adults and boys older than one year. One randomised trial had 70 patients, four cohort studies 400,000 patients, and seven case control studies 2,150 patients (Table 1).

The three types of study design had a similar result in that the relative risk was about 0.15, indicating a large reduction in UTI rates in circumcised over uncircumcised boys. There was consistency within the study types as well, and the four large cohort studies are shown in Figure 1. The absolute size of the effect varied, so that a number needed to treat calculated from the cohort studies was about 100, while that calculated from case-control studies was 6. The reason for the difference was the very large event rates in case control studies compared with cohorts; in uncircumcised boys the event rate was 20-fold higher in case control studies.

Figure 1: UTI rates in cohort studies, in circumcised and uncircumcised boys



A useful calculation was for benefit and harm for circumcision for UTI at different rates of urinary tract infection. Table 2 shows calculations in 1000 boys for three rates of UTI, a normal UTI rate of 1%, a higher rate of 10% in boys with recurrent UTI, and a highest rate of 30% in boys with vesicouretic reflux. Using the same 10-fold reduction with circumcision, and known information about complications at 2%, it is clear that the benefits of circumcision outweigh harm when UTI rates are higher than 2-3%.

Comment

This is a useful and informative systematic review. Apparently there have been differences of opinion on how big the effect of circumcision is on UTI, though the individual studies here showed a remarkable consistency. Circumcision substantially reduces the rate of UTI, and when UTI rates are 3% or more benefits probably outweigh harms. And there is a lesson, that while the statistical results may be the same between study designs, the absolute rates can vary substantially, with real consequences for calculations of NNT.

Reference:

1 D Singh-Grewal et al. Circumcision for the prevention of urinary tract infection in boys: a systematic review of randomised trials and observational studies. Ar-

Table 1: Effect of circumcision on UTI rates in studies of different design

Study type	Number of		UTI (%) in		Relative risk (95% CI)	NNT (95% CI)
	Studies	Patients	Circumcised	Uncircumcised		
Randomised trial	1	70	0.00	10	0.14 (0.01 to 2.6)	
Cohort	4	400700	0.13	1.1	0.13 (0.12 to 0.15)	99 (93 to 106)
Case control	7	2148	5.0	22	0.16 (0.11 to 0.23)	6 (5 to 7)

Table 2: Benefit and harm of circumcision at different UTI rates

Patient group	Risk of UTI (%)	UTI numbers in 1000 patients		Number of	
		Circumcised	Uncircumcised	UTI prevented	Complications of circumcision
Normal	1	1	10	9	20
Previous UTI	10	13	100	87	20
Vesicouretic reflux	30	39	300	261	20

Smoking and leg bypass grafts

Smoking is the single largest risk factor for development of peripheral arterial disease. Those with the worst disease necessitating bypass surgery have almost always been heavy smokers. The problem is to stop them smoking, not just because stopping smoking is a good thing of itself, but also because continued smoking is thought to affect the effectiveness of surgery. Meta-analysis of studies provides useful information about the effects of smoking, as well as a good example of how observational study design can affect results [1].

Meta-analysis

Authors searched for articles about peripheral artery disease from 1950 to 2004 in at least four electronic databases including Cochrane, together with reference lists and reviews. To be included studies had to provide information on primary, secondary, or cumulative patency rates of arterial reconstructive surgery in the lower extremities, and report rates for smokers and nonsmokers.

Patency could be determined by several methods, including pulse examination. Any form of determination of smoking was used. Any type of reconstructive surgery was allowed, with any graft material used in the surgery, with any length of follow up.

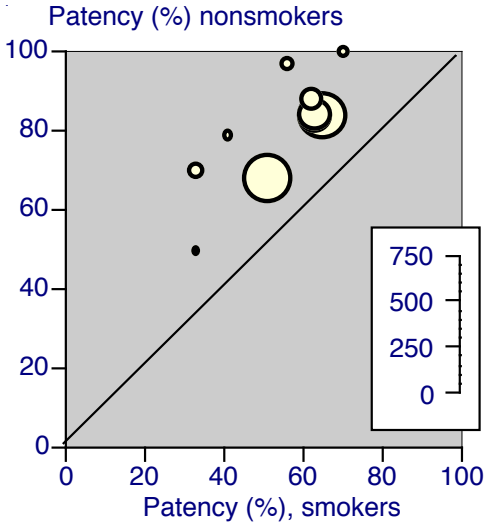
Results

Information was available from 29 studies, four randomised trials (870 patients), 12 prospective (1722 patients) and 13 retrospective (2894 patients) observational studies. Follow up varied between six months and 10 years, but most studies were between three and five years.

The main results are in Table 1. Prospective studies (including one randomised trial, Figure 1) produced a large difference between nonsmokers and smokers, with smoking having a number needed to harm for reduced patency of 4 (3.4 to 5.2). Retrospective studies also had a greater patency rate for nonsmokers over smokers, but the deleterious consequences of smoking were less, with a number needed to harm of 9 (6 to 16).

In prospective studies, how smoking was determined made no difference in patency rates, nor was there a difference between follow up times of less than two years and more than two years.

Figure 1: Individual results on graft patency in smokers and nonsmokers in prospective studies



Comment

So smoking is not a good idea if you have lower extremity bypass surgery. Yet over half of the patients in these studies smoked, and at least some of those who said they didn't smoke probably lied about it.

The difference between prospective and retrospective studies here was large and statistically significant. There was no overlap of NNH confidence intervals, and a statistical test showed a high level of difference ($p < 0.001$).

Why was this? There was no obvious answer in the trials, but it may just be some interaction between the type of study, prospective or retrospective, and the propensity of patients to fib about smoking. There wasn't much difference between the two types of study for graft patency isnsmokers, but a much lower patency in "nonsmokers" in retrospective studies. Of course, it may also be chance, but the numbers were quite large, with significant events, so this is an unlikely explanation.

So what can we tell patients about smoking and lower extremity bypass surgery? If they don't smoke, the chance of the surgery still working three to five years later is about 4 to 1 on. If they smoke, it's only evens, and they are likely to be in trouble.

Reference:

1 EM Willigendael et al. Smoking and the patency of lower extremity bypass grafts: a meta-analysis. Journal of Vascular Surgery 2005 42: 67-74.

Table 1: Effect of smoking on lower extremity bypass surgery grafts in smokers and nonsmokers, by study design

Study type	Number of		Patency (%) in		Relative risk (95% CI)	NNH (95% CI)
	Studies	Patients	Nonsmokers	Smokers		
Prospective	10	1198	81	57	1.4 (1.3 to 1.6)	4.1 (3.4 to 5.2)
Retrospective	7	1594	60	49	1.2 (1.1 to 1.4)	8.8 (6.2 to 16)
All	17	2792	69	53	1.4 (1.3 to 1.5)	6.3 (5.1 to 8.2)

Implementing better diabetes care

There is quite a lot of evidence about clinical interventions to improve diabetes care, as well as evidence about management interventions that can improve care and patient satisfaction. We know that improved care leads to better glycaemic control, and that, together with appropriate control of weight, cholesterol, and blood pressure, results in improved outcomes.

More problematic is the business of putting it all together in an overall package of care to deliver the goods. Here we move from randomised to observational studies, and from the hundreds and thousands to the tens of thousands or millions of patients. A report from Israel [1] demonstrates that putting it all together begins to deliver the goods.

Problems and changes

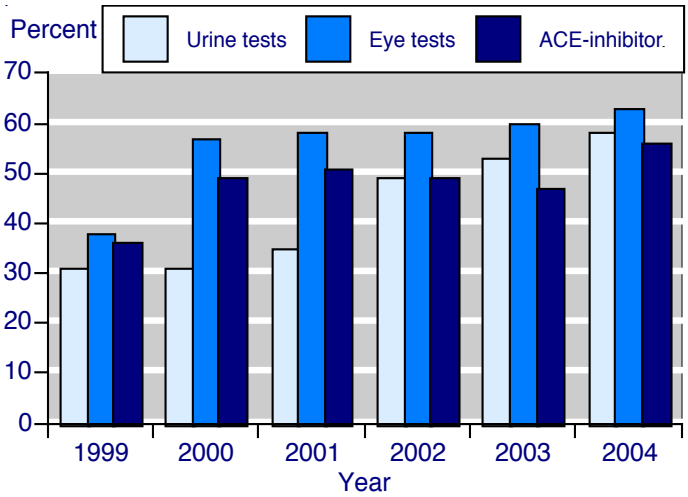
The setting was a healthcare services organisation serving a quarter of the population of Israel (about 1.5 million people), with almost 3% diagnosed as diabetic. Issues with structure of care provision, the problems caused, and the solutions used in a new care programme are shown in Table 1.

There were two main thrusts to the programme:

- One was a comprehensive diabetic registry created because all primary care physicians in the organisation use the same electronic medical record, maintained on a central computer.
- The other was a multidisciplinary protocol for care, based on best available evidence, and developed jointly between primary care and diabetic clinics.

Each clinical diabetologist was instructed to spend 20% of their working time on different aspects of disease management instead of direct patient care, time being created by discharging patients not on insulin to primary care. Education, feedback, patient support groups and the working of the registry were the core features of the programme.

Figure 1: Process indicators in implementing better diabetes care



The population was the 22,000 members continuously in the diabetes register from 1999 to 2004. The intervention began in January 2000.

Results

Urine collections for micro-albumin analysis, eye tests, and prescription of ACE-inhibitors increased with the new programme (Figure 1), as did testing of glycosylated haemoglobin and LDL-cholesterol from 74% in 1999 to 88% in 2004. Other process indicators, like blood pressure and BMI were not available.

Mean levels of glycosylated haemoglobin fell from 8.1% in 1999 to 7.8% in 2004. An absolute reduction of 2% or more occurred in 6% of patients, and of 1% or more in 16%.

Mean levels of LDL-cholesterol fell from 3.3 mmol/L in 1999 to 2.9 mmol/L in 2004. In 12,000 patients having LDL-cholesterol tested each year, the average absolute reduction was by 0.34 mmol/L.

The number of patients seen at diabetic clinics rose by 61% with only a 23% increase in work hours. Resources increased less than the increase in diabetic numbers.

Table 1: Problems and solutions around implementing better diabetes care

Structure issues	Problem	Solution
Self-referral to diabetic clinic	Overloaded clinics	GP referral only
	Imbalance between need and delivery of specialist treatment	Clinic outreach for specialist treatment by need, via GP
Absence of guidelines	No agreed standard of care	Multidisciplinary consensus programme development and dissemination
Absence of central monitoring	No information for decision-making	Automated registry provided feedback at all levels
	No feedback	
Split responsibility for care	Little communication between providers	Centralised monitoring
	Difficulties in coordinating care	Diabetic clinic responsible for patient care, coordination, and education, and for patient empowerment
	Specialists unwilling to discharge patients to GP	More patients with complex problems treated by specialists

Comment

The study followed a large cohort over five years, before and after introduction of a reformed programme of diabetic care. Measured indicators of the care process demonstrated large improvements, and measured indicators of benefit improved. Mean glycosylated haemoglobin levels dropped by 4% over five years, in contrast to the increase of 7% usually seen as patients age by five years.

Keys to success were good information technology, communal ownership of the programme, and being able to step back and make a big change. The diabetes register and common ownership maximised care and spotted when patients needed specialist care. Evidence, plus thoughtful management, provided better care, at a lower cost per patient. Impressive stuff.

Reference:

1 AD Heymann et al. The implementation of managed care for diabetes using medical informatics in a large preferred provider organisation. Diabetes Research and Clinical Practice 2005 (e-publication ahead of print).

Funnel plots: is seeing believing?

Publication bias is one of those chestnuts that is difficult to deal with. At its most nihilistic the argument is that negative studies are not published, so that any positive finding in a systematic review or meta-analysis will be balanced by many more studies showing no effect: nothing works. Invoking the spectre of publication bias is a useful support for doing nothing.

That is not to say that publication bias does not occur, because it surely does. The problem is knowing what you don't know. One solution has been the use of funnel plots, where some measure of effect of treatment (like odds ratio) is plotted against some measure of size or variance. The claim is that the absence of publication bias produces symmetry, and the presence of publication bias asymmetry.

Expert advice on systematic review and meta-analysis is to check for publication bias using funnel plots, both for writers and readers. Many journals insist on them, despite many publications demonstrating uncertainty and caveats, most of which seem to have gone unheeded. A new study [1] demonstrates the difficulty in recognising asymmetry in funnel plots.

Study

Researchers produced a number of funnel plots based on simulated meta-analyses, each with 10 studies, with sample size generated randomly for individual studies between 50 and 500, with a median of 158. This is typical of many meta-analyses. Publication bias was generated in some by allowing studies with lower P-values and higher sample size to be more likely to be included, and analyses could be statistically homogeneous or heterogeneous.

Twenty-two funnel plots were shown to 41 participants who were undertaking a meta-analysis course, and clinical researchers with some knowledge of meta-analysis. They were asked to grade funnel plots, using written instructions, as to whether there was evidence of publication bias (yes, no, maybe).

Results

About half (53% for all, 55% excluding maybe responses) the plots were correctly identified, with no difference between types of participant. That participants could not identify publication bias from funnel plots was not surprising, because a wide range of asymmetry was present in funnel plots with and without simulated publication bias, or with and without heterogeneity in 1,000 simulated meta-analyses (Table 1).

Comment

Basically then, most of us would feel uncomfortable about standing up and giving a quick lecture about funnel plots, and we couldn't spot publication bias in a funnel plot anyway. There is a large literature that confirms the finding in Table 1, that asymmetry exists with and without publication bias, so asymmetry tells us nothing about publication bias. Perhaps we ought to think it out again.

Reference:

1 N Terrin et al. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. Journal of Clinical Epidemiology 2005 58: 894-901.

Table 1: Heterogeneity and publication bias in simulated meta-analyses, and the degree of funnel plot asymmetry

		Asymmetry			
Heterogeneity	Publication bias	Very high	High	Moderate	Low
No	No	3	13	28	56
	Yes	17	18	32	34
Yes	No	4	14	31	51
	Yes	15	23	33	29

Raynaud's phenomenon update

Raynaud's phenomenon is an episodic event where the fingers, toes, ears, nose or jaw become pale and/or blue, often with pain, perhaps in response to cold or some other form of stress. About 1% or so of those with the phenomenon develop a connective tissue disease, most often some form of scleroderma. Studies of Raynaud's phenomenon are not common, even though as many as 1 in 30 to 1 in 10 adults may suffer at some time in their lives.

A Bandolier reader asked for an update on the latest data. Three studies caught Bandolier's attention: a new study on natural history in the community, an older meta-analysis about transition to secondary disease, and a new meta-analysis concerning treatment with calcium channel blockers.

Natural history [1]

The population used for this study was 1,525 participants in an ancillary study of children of original Framingham subjects. A standardised instrument for identifying Raynaud's phenomenon was administered in the early 1990s, and again about seven years later. Cases were identified as having Raynaud's phenomenon if in the last 12 months they met three or four of four criteria:

- unusual cold sensitivity
- unusual digit colour changes
- a positive response for blanching in comparison with a colour chart and response to a question about the palest their fingers ever get
- a positive response for blanching in comparison with actual photographs displaying digital blanching.

The study looked at prevalent Raynaud's phenomenon, and incident Raynaud's phenomenon defined as onset within the last 12 months. Those with Raynaud's phenomenon at baseline were further examined at seven years for persistence of the condition, or remission, where Raynaud's phenomenon had resolved.

Results

Baseline and follow up data were complete for 89% of the 1,525 participants, and the average follow up was 7.1 years. The average age was 54 years, with a range of 26 to 81 years. Prevalence was 11% for women and 8% for men, with incidence of about 2% a year (Table 1).

Table 1: Prevalence, incidence, and natural history of Raynaud's phenomenon in the community

	Women	Men
Number of subjects	717	641
Baseline prevalence (%)	11	7.8
Baseline incidence (%)	2.2	1.5
Number of cases	78	50
Persisting Raynaud's (%)	36	36
Remitted Raynaud's (%)	64	64

Table 2: Progression to other diseases among 639 people with Raynaud's phenomenon

Disease	Transitions	
	Number (total=81)	Percent of transitions
Systemic sclerosis	53	65
Mixed connective tissue disease	8	10
Sjögren syndrome	6	7
Rheumatoid arthritis	5	6
Systemic lupus erythematosus	4	5
Vasculitis	2	2
Myositis	2	2
Other	1	1

There was no relationship between Raynaud's phenomenon and age, BMI, occupational vibratory tool use, season of examination, place of residence, use of antihypertensive agent, or smoking. Two thirds of those with Raynaud's phenomenon at baseline did not fulfil criteria for Raynaud's phenomenon at follow up (Table 1).

Transition to other diseases [2]

This meta-analysis looked for articles that identified patients with Raynaud's phenomenon, excluded secondary causes, and followed up and re-evaluated patients.

Results

Ten studies included 639 patients. In them, the average age of onset of Raynaud's phenomenon was 34 years (range 23 to 46 years), and average age at entry into studies was 42 years, with Raynaud's phenomenon present for an average of 8 (range four to 21) years. Studies followed up patients with Raynaud's phenomenon for an average of 4.0 years, with 2,531 patient years of follow up.

During follow up, 81 of the 639 patients (13%) developed a secondary disease, 80 of which were connective tissue diseases (Table 2). Two thirds of these were systemic sclerosis. The average rate of transition measured from the onset of Raynaud's phenomenon was 1.4 (range 0.4 to 1.9) per 100 patient years, or 1% a year.

All studies measured one or more clinical or laboratory variables which potentially were predictors of clinical transition. Positive and negative predictive values calculated from studies are shown in Table 3. The best predictor of transition was abnormal capillary nailfold pattern, but the negative predictive value (Proportion of people with a negative test who are free of the target disorder) was high for several features.

Calcium channel blockers [3]

This review sought randomised trials in which calcium channel blockers were compared with placebo or other treatments, using standard search strategies, but searching only Medline. The main outcome was standardised to the number of attacks over a week.

Table 3: Predictive value of clinical and laboratory tests

Variable	Number of trials	Predictive value (%)	
		Positive	Negative
Abnormal nailfold capillaries	6	47	93
Cutaneous lesions	4	36	97
Positive antinuclear antibody test	9	30	93
Abnormal pulmonary function	4	39	88
Oesophageal dysmotility	4	33	88
Digital ulcer or pits	5	26	75

Results

Eighteen randomised, double blind, trials with placebo control were included, in which the mean frequency of attacks per week with control was 11 (95% CI 8 to 14). Almost all were crossover trials, lasting one to 10 weeks. Nifedipine at various doses and regimens featured in 13 of the studies. Trial size varied between six and 138 patients. Seven trials had treatment group sizes below 10 patients, and 13 group sizes below 25 patients.

Treatment with calcium channel blockers reduced attacks by five per week, but the largest effects were in the smaller trials. Only four trials had more than 25 patients per group; none showed calcium channel blocker to be better than placebo. These four larger trials reduced attacks by about 1 per week or less. While severity of attacks was also reduced overall, the larger trials showed no effect.

Comment

Sufferers from Raynaud's phenomenon can be given the good news that two thirds are likely to have a spontaneous resolution of symptoms in the future. The bad news is that some of them, 10%, will progress to a more serious condition.

Nor is there much good news on treatment. Only a single meta-analysis could be found, which, despite its headline claims, lent no real weight to the efficacy of calcium channel blockers. Nor did a search for randomised trials reveal any good large trials of sensible interventions with large beneficial effects published in recent years.

References:

1 LG Suter et al. The incidence and natural history of Raynaud's phenomenon in the community. Arthritis & Rheumatism 2005 52: 1259-1263.
2 G Spencer-Green. Outcomes in primary Raynaud phenomenon. A meta-analysis of the frequency, rates, and predictors of transition to secondary diseases. Archives of Internal Medicine 1998 158: 595-600.
3 AE Thompson, JE Pope. Calcium channel blockers for primary Raynaud's phenomenon: a meta-analysis. Rheumatology 2005 44: 145-150.

Quality control in systematic reviews

Bandolier 138 reported a Cochrane disaster, where a review about electronic information for patients reached completely the wrong answer through mistakes. The internal quality control systems of peer review in Cochrane failed to spot it. Well mistakes happen, sometimes large, but more often small. It is good to know that systematic quality control happens in Cochrane, demonstrating no major problems [1].

This study retrospectively repeated the data extraction in all systematic reviews published from the Cochrane Cystic Fibrosis and Genetic Disorders Group. The extraction was done by a statistician experienced in systematic reviews.

Results

There were 42 reviews, of which eight had found no studies to include, so 34 reviews formed the basis of the study. Fifty errors were found in 20 reviews, with 21 different error types. The most common errors were inconsistency between data calculation and results reported in text, and inconsistent reporting of significant versus nonsignificant results. While the errors made for changes to summary results, none altered the conclusions.

Comment

First of all, three cheers for the Cochrane Cystic Fibrosis and Genetic Disorders Group for having the wherewithal to do this quality control exercise, no mean feat. It demanded lots of hard work, and a certain amount of courage. Errors occurred commonly, but made little difference. Peer review by external referees did not spot them, despite four reviewers for each review (consumer, methodologist, statistician, subject expert). Processes have been changed to respond to these findings.

Errors are not limited to Cochrane reviews, or just to reviews. They are common, and perhaps as many as 1 in 10 papers has at least one major error. Minor errors are legion. Readers have to act as their own quality control, and not accept all written words as absolutely and unquestionably true. Question more.

Reference:

1 AP Jones et al. High prevalence but low impact of data extraction and reporting errors were found in Cochrane systematic reviews. Journal of Clinical Epidemiology 2005 58: 741-742.

EDITORS:
Andrew Moore Henry McQuay
Pain Relief Unit
The Churchill, Oxford OX3 7LJ
Editorial office: 01865 226132
Editorial fax: 01865 226978
Email: andrew.moore@pru.ox.ac.uk
Internet: www.ebandolier.com
ISSN 1353-9906